GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: January 3, 2002, 19:16:29 ; Search time 2882.41 Seconds
(without alignments)
18945.968 Million cell updates/sec

Title: US-09-497-822A-18
Perfect score: 5082
Sequence: 1 gagctctggacaaaattgag.....acaagcaaacaaaaaaaaaa 5082

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_htc:*
10: gb_est1:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	571.8	11.3	654	10	AW951855	AW951855 EST363925
c 2	501.4	9.9	568	10	AA524966	AA524966 nh35c11.s
3	462.2	9.1	724	10	AI317423	AI317423 uj22e09.y
4	380.6	7.5	388	10	AA229063	AA229063 nc49g07.s
c 5	369	7.3	369	10	AA229062	AA229062 nc49g07.r
6	361	7.1	413	11	T28396	T28396 EST41738 Hu
7	358.4	7.1	573	10	AW619116	AW619116 436 MARC
c 8	357.8	7.0	637	10	AI785019	AI785019 uj22e09.x
9	347.4	6.8	349	10	AA230070	AA230070 nc49b06.s
c 10	346	6.8	379	10	AA659567	AA659567 nt63a03.s
c 11	319	6.3	536	10	AI315085	AI315085 uj23g04.x
12	293.2	5.8	498	13	AZ017963	AZ017963 RPCI-23-2
c 13	285.4	5.6	292	10	AA229714	AA229714 nc49b06.r
14	247.6	4.9	395	13	AZ267276	AZ267276 RPCI-23-1
c 15	238	4.7	247	10	AI659563	AI659563 tu12d12.x
16	234.2	4.6	396	10	AW619096	AW619096 424 MARC
17	216.8	4.3	667	11	BI183393	BI183393 UNL-P-FN-
18	214.6	4.2	413	10	AI326670	AI326670 mo60g04.y
c 19	203.4	4.0	378	10	AI893550	AI893550 mo60g04.x
20	196.2	3.9	322	10	AI317523	AI317523 uj23g04.y
21	193.2	3.8	339	10	AW619117	AW619117 437 MARC
22	190.6	3.8	389	10	AA089319	AA089319 mo60g04.r
23	184.6	3.6	798	10	AU136010	AU136010 AU136010
24	174.6	3.4	613	11	BG938686	BG938686 cn27g02.x
25	173	3.4	643	11	BG345621	BG345621 dd89g10.y
26	171.8	3.4	621	11	BG886115	BG886115 dad73b08.
27	171.6	3.4	627	11	BG817008	BG817008 dad68h09.
28	167.6	3.3	293	10	AA561201	AA561201 vl28g06.r
29	164.6	3.2	546	11	BG997346	BG997346 MR4-HT105
30	163.6	3.2	720	10	AU139958	AU139958 AU139958
31	162.2	3.2	831	11	BG614664	BG614664 602642157
32	159.4	3.1	521	10	AW910225	AW910225 ur79c10.y
c 33	158.4	3.1	464	13	AZ303301	AZ303301 UP_479-1J
34	157.6	3.1	459	10	AA252234	AA252234 zr63h08.r
c 35	157.4	3.1	531	10	AW050106	AW050106 UI-M-BH1-
36	155.4	3.1	530	10	BE667525	BE667525 151944 MA
c 37	152.8	3.0	267	13	AZ939124	AZ939124 2M0197P20
c 38	147.6	2.9	683	10	AI829064	AI829064 wj38b09.x
39	147.6	2.9	868	11	BF542009	BF542009 602069152
c 40	146.8	2.9	601	10	AW270095	AW270095 xv37h09.x
c 41	146.8	2.9	691	10	AI828886	AI828886 wj37b09.x
42	146.8	2.9	929	10	BE572814	BE572814 601327869
43	146.6	2.9	215	13	AZ884598	AZ884598 RPCI-23-1
c 44	145.4	2.9	315	13	AZ936934	AZ936934 2M0193L19
45	142.8	2.8	216	10	BB588775	BB588775 BB588775

ALIGNMENTS

RESULT 1
AW951855
LOCUS AW951855 654 bp mRNA EST 01-JUN-2000
DEFINITION EST363925 MAGE resequences, MAGB Homo sapiens cDNA, mRNA sequence.
ACCESSION AW951855
VERSION AW951855.1 GI:8141534
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 654)
AUTHORS Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and Quackenbush,J.
TITLE Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished (2000)
COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@tigr.org
Plate: 32
Seq primer: Reverse.
FEATURES Location/Qualifiers
source 1. .654
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGB"
/note="Vector: pBluescriptSKm"
BASE COUNT 166 a 183 c 133 g 172 t
ORIGIN

Query Match 11.3%; Score 571.8; DB 10; Length 654;
Best Local Similarity 95.3%; Pred. No. 3.8e-63;
Matches 622; Conservative 0; Mismatches 27; Indels 4; Gaps 3;

```
Qy   4048 ttccgcaacttacacgtggacgaccagatggctgtcattcagtactcctggatggggctc 4107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTCCGCAACTTACACGTGGACGACCAGATGGCTGTCATTCAGTACTCCTGGATGGGGCTC 60

Qy   4108 atggtgtttgccatgggctggcgatccttcaccaatgtcaactccaggatgctctacttc 4167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATGGTGTTTGCCATGGGCTGGCGATCCTTCACCAATGTCAACTCCAGGATGCTCTACTTC 120

Qy   4168 gcccctgatctggtttttcaatgagtaccgcatgcacaagtcccggatgtacagccagtgt 4227
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCCCCTGATCTGGTTTTCAATGAGTACCGCATGCACAAGTCCCGGATGTACAGCCAGTGT 180

Qy   4228 gtccgaatgaggcacctctctcaagagtttggatggctccaaatcacccccaggaattc 4287
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTCCGAATGAGGCACCTCTCTCAAGAGTTTGGATGGCTCCAAATCACCCCCCAGGAATTC 240

Qy   4288 ctgtgcatgaaagcactgctactcttcagcattattccagtggatgggctgaaaaatcaa 4347
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CTGTGCATGAAAGCACTGCTACTCTTCAGCATTATTCCAGTGGATGGGCTGAAAAATCAA 300

Qy   4348 aaattctttgatgaacttcgaatgaactacatcaaggaactcgatcgtatcattgcatgc 4407
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AAATTCTTTGATGAACTTCGAATGAACTACATCAAGGAACTCGATCGTATCATTGCATGC 360

Qy   4408 aaaagaaaaaatcccacatcctgctcaagacgcttctaccagctcaccaagctcctggac 4467
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AAAAGAAAAAATCCCACATCCTGCTCAAGACGCTTCTACCAGCTCACCAAGCTCCTGGAC 420

Qy   4468 tccgtgcagcctattgcgagagagctgcatcagttcacttttgacctgctaatcaagtca 4527
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCCGTGCAGCCTATTGCGAGAGAGCTGCATCAGTTCACTTTTGACCTGCTAATCAAGTCA 480

Qy   4528 cacatggtgagcgtggactttccggaaatgatggcagagatcatctctgtgcaagtgccc 4587
          ||||||||||||||||||| |||||||||||||||| ||||||||||||||||||||||
Db    481 CACATGGTGAGCGTGGACTTTTCGGAAATGATGGCAGAAATCATCTCTGTGCAAGTGCCC 540

Qy   4588 aagatcctttctggg-aaagtcaagcccatctattt-ccacacccagtgaagcattggaa 4645
          |||||||||| |||| |||||||||||| ||| ||| |||||||||||||| | ||||||
Db    541 AAGATCCTTTTTGGGAAAAGTCAAGCCCTTCTTTTTCCCACACCCAGTGAAACCTTGGAA 600

Qy   4646 a--ccctatttccccaccccagctcatgccccctttcagatgtcttctgcctg 4696
          |  |||| ||||||||||||   |     ||||  |  ||     |||  |||
Db    601 AACCCCTTTTTCCCCACCCCCATTTTATGCCCCCCTTTGAAAAAGTCTTTCTG 653
```

RESULT 2
AA524966/c
LOCUS AA524966 568 bp mRNA EST 05-AUG-1997
DEFINITION nh35c11.s1 NCI_CGAP_Pr3 Homo sapiens cDNA clone IMAGE:954356 similar to gb:M23263 ANDROGEN RECEPTOR (HUMAN);, mRNA sequence.
ACCESSION AA524966
VERSION AA524966.1 GI:2265894
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 568)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui, M.D., Michael Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 362 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 383.
FEATURES Location/Qualifiers
source 1. .568
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:954356"
/clone_lib="NCI_CGAP_Pr3"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/note="Vector: pAMP10; Site_1: Not1; Site_2: EcoRI; 1st strand cDNA was primed with oligo(dT)17 on 50 ng of DNAse-treated, total cellular RNA obtained from 5,000-10,000 microdissected cells histologically-determined to be fully malignant prostate cancer cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into pAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."
BASE COUNT 186 a 99 c 153 g 130 t
ORIGIN

Query Match 9.9%; Score 501.4; DB 10; Length 568;
Best Local Similarity 98.5%; Pred. No. 2.9e-54;
Matches 527; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

```
Qy   4529 acatggtgagcgtggacttttccggaaatgatggcagagatcatctctgtgcaagtg-ccc 4587
          |||||||||||||||   |||||| |||||||| |||||||||||||||||||||| ||| |
Db    550 ACATGGTGAGCGTGACTTTTCCGAAAATGATGCCAGAGATCATCTCTGTGCAAGTGCCCC 491

Qy   4588 aagatcctttctgggaaagtcaagcccatctatttccacacccagtgaagcattggaaac 4647
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db    490 AAGATCCTTTCTGGGAAAGTCAAGCCCATCTATTTCCACACCCAGTGAAGCA-TGGAAAC 432

Qy   4648 cctatttccccaccccagctcatgccccctttcagatgtcttctgcctgttataactctg 4707
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    431 CCTATTTCCCCACCCCAGCTCATGCCCCCTTTCAGATGTCTTCTGCCTGTTATAACTCTG 372

Qy   4708 cactactcctctgcagtgccttggggaatttcctctattgatgtacagtctgtcatgaac 4767
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    371 CACTACTCCTCTGCAGTGCCTTGGGGAATTTCCTCTATTGATGTACAGTCTGTCATGAAC 312

Qy   4768 atgttcctgaatttctatttgctgggctttttttttctctttctctcctttctttttcttc 4827
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    311 ATGTTCCTGAATTCTATTTGCTGGGCTTTTTTTTTCTCTTTCTCTCCTTTCTTTTTCTTC 252

Qy   4828 ttccctccctatctaaccctcccatggcaccttcagactttgcttcccattgtggctcct 4887
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    251 TTCCCTCCCTATCTAACCCTCCCATGGCACCTTCAGACTTTGCTTCCCATTGTGGCTCCT 192

Qy   4888 atctgtgttttgaatggtgttgtatgcctttaaatctgtgatgatcctcatatggcccag 4947
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    191 ATCTGTGTTTTGAATGGTGTTGTATGCCTTTAAATCTGTGATGATCCTCATATGGCCCAG 132

Qy   4948 tgtcaagttgtgcttgtttacagcactactctgtgccagccacacaaacgtttacttatc 5007
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    131 TGTCAAGTTGTGCTTGTTTACAGCACTACTCTGTGCCAGCCACACAAACGTTTACTTATC 72

Qy   5008 ttatgccacgggaagtttagagagctaagattatctggggaaatcaaaacaaaaa 5062
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     71 TTATGCCACGGGAAGTTTAGAGAGCTAAGATTATCTGAGGAAATCAAAACAAAAA 17
```

RESULT 3

```
AI317423
LOCUS       AI317423      724 bp    mRNA            EST       17-DEC-1998
DEFINITION  uj22e09.y1 Sugano mouse kidney mkia Mus musculus cDNA clone
            IMAGE:1920712 5' similar to gb:X53779 Mouse mRNA for androgen
            receptor (MOUSE);, mRNA sequence.
ACCESSION   AI317423
VERSION     AI317423.1  GI:4032690
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 724)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:977004
            Seq primer: custom primer used
            High quality sequence stop: 515.
FEATURES             Location/Qualifiers
     source          1. .724
                     /organism="Mus musculus"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1920712"
                     /clone_lib="Sugano mouse kidney mkia"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pME18S-FL3; Site_1: DraIII
                     (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a  DraIII adaptor [TGTTGGCCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert.  Size selection was
                     performed to exclude fragments <1.5kb.  Library
                     constructed by Dr. Sumio Sugano (University of Tokyo
                     Institute of Medical Science).  Custom primers  for
                     sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end
                     primer CGACCTGCAGCTCGAGCACA."
BASE COUNT      190 a    181 c    192 g    161 t
ORIGIN
```

```
  Query Match             9.1%;  Score 462.2;  DB 10;  Length 724;
  Best Local Similarity   82.8%;  Pred. No. 2.4e-49;
  Matches  564;  Conservative    0;  Mismatches  113;  Indels    4;  Gaps    3;

Qy   3257 gcggcggcggcggcggcggcggcggcggcggcggcgaggcgggagctgtagccccctacg 3316
          ||   ||| |||||||| || ||   | ||     | | |  | |  ||   |||||| |
Db     16 GCCTGGGCTGCGGCGGCAGC-GCAATGCCGCTATGGGGACTTGGGTAGTCAGCCCCTATG 74

Qy   3317 gctacactcggccccctcaggggctggcgggccaggaaagcgacttcaccgcacctgatg 3376
          |||||||||||||||||||||||||| |  ||||||| || |||| | | ||  | || |
Db     75 GCTACACTCGGCCCCCTCAGGGGCTGACAAGCCAGGAGAGTGACTACTCTGCCTCCGAAG 134

Qy   3377 tgtggtaccctggcggcatggtgagcagagtgccctatcccagtcccacttgtgtcaaaa 3436
          ||||||| ||||| ||  | |||| |||||| |||||||||||||||| ||||||||||
```

```
Db    135 TGTGGTATCCTGGTGGAGTTGTGAACAGAGTACCCTATCCCAGTCCCAATTGTGTCAAAA 194

Qy   3437 gcgaaatgggcccctggatggatagctactccggaccttacggggacatgcgtttggaga 3496
          | |||||||| || |||||||| | |||||||||||| || ||||||| |||||||| |
Db    195 GTGAAATGGGACCTTGGATGGAGAACTACTCCGGACCCTATGGGGACACGCGTTTGGACA 254

Qy   3497 ctgccagggaccatgttttgcccattgactattactttccaccccagaagacctgcctga 3556
          | |||||||||||||||| ||||| ||||||||||||||||||||||||||||||||||
Db    255 GTACCAGGGACCATGTTTTACCCATCGACTATTACTTTCCACCCCAGAAGACCTGCCTGA 314

Qy   3557 tctgtggagatgaagcttctgggtgtcactatggagctctcacatgtggaagctgcaagg 3616
          ||||||||||||||||||||||| |||||||| |||||||||| ||||| ||||||||||
Db    315 TCTGTGGAGATGAAGCTTCTGGCTGTCACTACGGAGCTCTCACTTGTGGCAGCTGCAAGG 374

Qy   3617 tcttcttcaaaagagccgctgaagggaaacagaagtacctgtgcgccagcagaaatgatt 3676
          |||||||||||||||||||||||||||||||||||| || || |||||||||| ||||
Db    375 TCTTCTTCAAAAGAGCCGCTGAAGGGAAACAGAAGTATCTATGTGCCAGCAGAAACGATT 434

Qy   3677 gcactattgataaattccgaaggaaaaattgtccatcttgtcgtcttcggaaatgttatg 3736
          | || |||||||||| || |||||||||| |||||||||||||| ||||||||||||
Db    435 GTACCATTGATAAATTTCGGAGGAAAAATTGCCCATCTTGTCGTCTCCGGAAATGTTATG 494

Qy   3737 aagcagggatgactctgggagcccggaagctgaagaaacttggtaatctgaaactacagg 3796
          ||||| |||||||||||||||| || |||||||||||||||| ||||| ||||||||||
Db    495 AAGCAAGGATGACTCTGGGAGCTCGTAAGCTGAAGAAACTTGGAAATCTAAAACTACAGG 554

Qy   3797 aggaaggagaggcttccagcaccaccagccccactgaggagacaacccagaagctgacag 3856
          ||||||||||    ||||   |  ||||||||||||||| || |||||||| |||| |
Db    555 AGGAAGGAGAAAACTCCAATGCTGGCAGCCCCACTGAGGACCCATCCCAGAAGATGACTG 614

Qy   3857 tgtcacacattgaaggc-tatgaatgtcagcccatcttt--ctgaatgtcctggaagcca 3913
          | |||||||||||| || | |||||||||||| ||||||   | |    ||||||||||
Db    615 TATCACACATTGAAAGCTTTTGAATGTCAGCCTATCTTTTCTTAACGTCCCTGGAAGCCA 674

Qy   3914 ttgagccaggtgtagtgtgtg 3934
          ||    |    | |  |||||
Db    675 TTTGAGCCCAGGAATGGTGTG 695
```

```
RESULT   4
AA229063
LOCUS       AA229063      388 bp    mRNA            EST       21-AUG-1997
DEFINITION  nc49g07.s1 NCI_CGAP_Pr3 Homo sapiens cDNA clone IMAGE:1011516
            similar to gb:M23263 ANDROGEN RECEPTOR (HUMAN);, mRNA sequence.
ACCESSION   AA229063
VERSION     AA229063.1  GI:1852047
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 388)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui, M.D.
            , Michael Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: David B. Krizman, Ph.D.
            cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -41m13 fwd. ET from Amersham
            High quality sequence stop: 366.
FEATURES             Location/Qualifiers
     source          1. .388
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1011516"
```

/clone_lib="NCI_CGAP_Pr3"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/note="Vector: pAMP10; Site_1: Not1; Site_2: EcoRI; 1st strand cDNA was primed with oligo(dT)17 on 50 ng of DNAse-treated, total cellular RNA obtained from 5,000-10,000 microdissected cells histologically-determined to be fully malignant prostate cancer cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into pAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."
BASE COUNT 77 a 113 c 72 g 126 t
ORIGIN

Query Match 7.5%; Score 380.6; DB 10; Length 388;
Best Local Similarity 99.0%; Pred. No. 5.6e-39;
Matches 383; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy   4442 tctaccagctcaccaagctcctggactccgtgcagcctattgcgagagagctgcatcagt 4501
          || |||||||||||||||||||||||||  | ||||||||||||||||||||||||||
Db      2 TCGACCAGCTCACCAAGCTCCTGGACTCCGCTCCGCCTATTGCGAGAGAGCTGCATCAGT 61

Qy   4502 tcacttttgacctgctaatcaagtcacacatggtgagcgtggactttccggaaatgatgg 4561
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     62 TCACTTTTGACCTGCTAATCAAGTCACACATGGTGAGCGTGGACTTTCCGGAAATGATGG 121

Qy   4562 cagagatcatctctgtgcaagtgcccaagatcctttctgggaaagtcaagcccatctatt 4621
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    122 CAGAGATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTATT 181

Qy   4622 tccacacccagtgaagcattggaaaccctatttccccaccccagctcatgccccctttca 4681
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    182 TCCACACCCAGTGAAGCATTGGAAACCCTATTTCCCCACCCCAGCTCATGCCCCCTTTCA 241

Qy   4682 gatgtcttctgcctgttataactctgcactactcctctgcagtgccttggggaatttcct 4741
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    242 GATGTCTTCTGCCTGTTATAACTCTGCACTACTCCTCTGCAGTGCCTTGGGGAATTTCCT 301

Qy   4742 ctattgatgtacagtctgtcatgaacatgttcctgaattctatttgctgggctttttttt 4801
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    302 CTATTGATGTACAGTCTGTCATGAACATGTTCCTGAATTCTATTTGCTGGGCTTTTTTTT 361

Qy   4802 tctctttctctcctttctttttcttct 4828
          |||||||||||||||||||||||||||
Db    362 TCTCTTTCTCTCCTTTCTTTTTCTTCT 388
```

RESULT 5
AA229062/c
LOCUS AA229062 369 bp mRNA EST 21-AUG-1997
DEFINITION nc49g07.r1 NCI_CGAP_Pr3 Homo sapiens cDNA clone IMAGE:1011516 similar to gb:M23263 ANDROGEN RECEPTOR (HUMAN);, mRNA sequence.
ACCESSION AA229062
VERSION AA229062.1 GI:1852046
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 369)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui, M.D., Michael Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 341.
FEATURES Location/Qualifiers
source 1. .369
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1011516"
/clone_lib="NCI_CGAP_Pr3"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/note="Vector: pAMP10; Site_1: Not1; Site_2: EcoRI; 1st strand cDNA was primed with oligo(dT)17 on 50 ng of DNAse-treated, total cellular RNA obtained from 5,000-10,000 microdissected cells histologically-determined to be fully malignant prostate cancer cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into pAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."
BASE COUNT 142 a 67 c 93 g 67 t
ORIGIN

Query Match 7.3%; Score 369; DB 10; Length 369;
Best Local Similarity 100.0%; Pred. No. 1.6e-37;
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy   4680 cagatgtcttctgcctgttataactctgcactactcctctgcagtgccttggggaatttc 4739
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    369 CAGATGTCTTCTGCCTGTTATAACTCTGCACTACTCCTCTGCAGTGCCTTGGGGAATTTC 310

Qy   4740 ctctattgatgtacagtctgtcatgaacatgttcctgaattctatttgctgggctttttt 4799
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    309 CTCTATTGATGTACAGTCTGTCATGAACATGTTCCTGAATTCTATTTGCTGGGCTTTTTT 250

Qy   4800 tttctctttctctccttttcttttttcttcttccctccctatctaaccctcccatggcacct 4859
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    249 TTTCTCTTTCTCTCCTTTCTTTTTCTTCTTCCCTCCCTATCTAACCCTCCCATGGCACCT 190

Qy   4860 tcagactttgcttcccattgtggctcctatctgtgttttgaatggtgttgtatgccttta 4919
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    189 TCAGACTTTGCTTCCCATTGTGGCTCCTATCTGTGTTTTGAATGGTGTTGTATGCCTTTA 130

Qy   4920 aatctgtgatgatcctcatatggcccagtgtcaagttgtgcttgtttacagcactactct 4979
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    129 AATCTGTGATGATCCTCATATGGCCCAGTGTCAAGTTGTGCTTGTTTACAGCACTACTCT 70

Qy   4980 gtgccagccacacaaacgtttacttatcttatgccacgggaagtttagagagctaagatt 5039
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     69 GTGCCAGCCACACAAACGTTTACTTATCTTATGCCACGGGAAGTTTAGAGAGCTAAGATT 10

Qy   5040 atctgggga 5048
          |||||||||
Db      9 ATCTGGGGA 1
```

RESULT 6
T28396
LOCUS T28396 413 bp mRNA EST 06-SEP-1995
DEFINITION EST41738 Human Uterus Homo sapiens cDNA 5' end similar to androgen receptor (HT:1334), mRNA sequence.
ACCESSION T28396
VERSION T28396.1 GI:610494
KEYWORDS EST.

```
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 413)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
            ,C.J., Lee,N., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
            ,O., Sutton,G., Blake,J.A., Brandon,R.C., Chiu,M.-W., Clayton,R.A.,
            Cline,R.T., Cotton,M.D., Earle-Hughes,J., Fine,L.D., FitzGerald
            ,L.M., FitzHugh,W.M., Fritchman,J.L., Geoghagen,N.S.M., Glodek,A.,
            Gnehm,C.L., Hanna,M.C., Hedblom,E., HinkleJr,P.S., Kelley,J.M.,
            Klimek,K.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pellegrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.-J.,
            Dimke,D., Feng,P., Ferrie,A., Fischer,C., Hastings,G.A., He,W.-W.,
            Hu,J.-S., Greene,J.M., Gruber,J., Hudson,P., Kim,A., Kozak,D.L.,
            Kunsch,C., Ji,H., Li,H., Meissner,P.S., Olsen,H., Raymond,L., Wei
            ,Y.-F., Wing,J., Xu,C., Yu,G.-L., Ruben,S.M., Dillon,P.J., Fannon
            ,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
            Venter,J.C.
  TITLE     Initial Assessment of Human Gene Diversity and Expression Patterns
            Based Upon 83 Million Basepairs of cDNA Sequence
  JOURNAL   Nature 377, 3-174 (1995)
  MEDLINE   96026280
COMMENT     Contact: Venter, JC
            The Institute for Genomic Research
            932 Clopper Rd, Gaithersburg, MD 20878
            Tel: 3018699056
            Fax: 3018699423
            Email: tdbinfo@tdb.tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please contact the TIGR Database
            (tdbinfo@tdb.tigr.org)
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .413
                     /organism="Homo sapiens"
                     /db_xref="ATCC (inhost):102309"
                     /db_xref="taxon:9606"
                     /clone_lib="Human Uterus"
                     /note="Organ: uterus"
BASE COUNT      100 a    108 c     92 g    107 t      6 others
ORIGIN
```

```
  Query Match             7.1%;  Score 361;  DB 11;  Length 413;
  Best Local Similarity   94.5%;  Pred. No. 1.6e-36;
  Matches  381;  Conservative    0;  Mismatches   21;  Indels    1;  Gaps    1;

Qy   4047 cttccgcaacttacacgtggacgaccagatggctgtcattcagtactcctggatggggct 4106
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db      1 CTTCCGCAACTTACACGTGGACGACCAGATGGCTGTNATTCAGTACTCCTGGATGGGGCT 60

Qy   4107 catggtgtttgccatgggctggcgatccttcaccaatgtcaactccaggatgctctactt 4166
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     61 CATGGTGTTTGCCATGGGCTGGCGATCCTTCACCAATGTNAACTCCAGGATGCTCTACTT 120

Qy   4167 cgcccctgatctggttttcaatgagtaccgcatgcacaagtcccggatgtacagccagtg 4226
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CGCCCCTNATCTGGTTTTCAATGAGTACCGCATGCACAAGTCCCGGATGTACAGCCAGTG 180

Qy   4227 tgtccgaatgaggcacctctctctcaagagtttggatggctccaaatcacccccccaggaatt 4286
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TGTCCGAATGAGGCACCTCTCTCAAGAGTTTGGATGGCTCCAAATCACCCCCCAGGAATT 240

Qy   4287 cctgtgcatgaaagcactgctactcttcagcattattccagtggatgggctgaaaaatca 4346
          |||||||||||||||||||||||||||||||||||||||||||    ||    |||||||
Db    241 CCTGTGCATGAAAGCACTGCTACTCTTCAGCATTATTCCAGTGGGATGGGCTGAAAATCA 300

Qy   4347 aaaattctttgatgaacttcgaatgaactacatcaaggaactcgatcgtatcattgcatg 4406
          |||||||||||||| |||||| ||| |||||| ||||| ||||| ||||||||||||||||
```

```
Db    301 AAAATTCTTTGATGNACTTCGGATGNACTACA-CAAGGGACTCGGTCGTATCATTGCATG 359

Qy   4407 caaaagaaaaaatcccacatcctgctcaagacgcttctaccag 4449
          |||||| |||||||||||||||||| |  ||||| || ||||||
Db    360 CAAAAGGAAAAATCCCACATCCTGNTTCAGACGGTTTTACCAG 402
```

```
RESULT   7
AW619116
LOCUS       AW619116      573 bp    mRNA            EST       24-MAR-2000
DEFINITION  436 MARC PBE Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   AW619116
VERSION     AW619116.1  GI:7325300
KEYWORDS    EST.
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 573)
  AUTHORS   Smith,T.P.L., Fahrenkrug,S.C., Rohrer,G.A., Simmen,F.A., Rexroad
            ,C.E. and Keele,J.W.
  TITLE     Mapping of expressed sequence tags from a porcine early embryonic
            cDNA library
  JOURNAL   Anim. Genet. 32 (2), 66-72 (2001)
  MEDLINE   21314990
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: GGAAACAGCTATGACCATG
            BACKWARD: GTAAAACGACGGCCAGT
            Seq primer: AATTAACCCTCACTAAAGGG.
FEATURES             Location/Qualifiers
     source          1. .573
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
                     /clone_lib="MARC PBE"
                     /tissue_type="Day 12 whole embryos"
                     /lab_host="XLOLR"
                     /note="Vector: pBLUESCRIPT SK-; Site_1: EcoRI; Site_2:
                     XhoI; Library made from pool of embryos in spherical and
                     filamentous stages of development (7.5% and 92.5%,
                     respectively, of each stage) as described in Choi et al,
                     Endocrinology 137, 1457-67, 1996."
BASE COUNT      125 a    199 c    162 g     87 t
ORIGIN
```

```
  Query Match             7.1%;  Score 358.4;  DB 10;  Length 573;
  Best Local Similarity   77.7%;  Pred. No. 3.1e-36;
  Matches  471;  Conservative    0;  Mismatches   81;  Indels   54;  Gaps    1;

Qy   1863 gatggaagtgcagttagggctgggaagggtctaccctcggccgccgtccaagacctaccg 1922
          | |||||||||||||||||||||| ||||||||||    ||||| |||||||||  ||
Db     10 GTTGGAAGTGCAGTTAGGGCTGGGGAGGGTCTACCCCTTTCCGCCTTCCAAGACCTTTCG 69

Qy   1923 aggagctttccagaatctgttccagagcgtgcgcgaagtgatccagaacccgggccccag 1982
          ||| |||||||||| |||||||| || || |||||||||||||||||||| || ||| |
Db     70 AGGTGCTTTCCAGAACCTGTTCCAAAGTGTACGCGAAGTGATCCAGAACCCAGGTCCCCG 129

Qy   1983 gcacccagaggccgcgagcgcagcacctcccggcgccagtttgctgctgctgcagcagca 2042
          |||||| |||||||||||||||||||||||||||||||| |||||| || || ||||| |||
Db    130 ACACCCTGAGGCCGCGAGCGCAGCACCTCCCGGCGCCCGTTTGCAGCAGCAGCAGCTGCA 189

Qy   2043 gcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagca 2102
          ||||||| ||       |  | || |||
Db    190 GCAGCAGGAGACCAGTCCCCGGCGGCA---------------------------------- 216
```

```
Qy   2103 gcagcaagagactagccccaggcagcagcagcagcagcagggtgaggatggttctcccca 2162
                              |||||||||||||||| | |||||||||| ||||||||
Db    217 ---------------------GCAGCAGCAGCAGCCGAGTGAGGATGGCTCTCCCCA 255

Qy   2163 agcccatcgtagaggccccacaggctacctggtcctggatgaggaacagcaaccttcaca 2222
          || |||  | |||||||||||||||||||||||| || ||| ||| ||||||| ||||||||
Db    256 AGTCCAGAGCAGAGGCCCCACAGGCTACCTGGCCCCGGACGAGAAACAGCAGCCTTCACA 315

Qy   2223 gccgcagtcggccctggagtgccaccccgagagaggttgcgtcccagagcctggagccgc 2282
           | ||| || ||||  |||||||||| ||||| || |||  ||||||||| ||||| ||
Db    316 ACAGCAATCAGCCCCTGAGTGCCACCCGGAGAGCGGCTGCACCCCAGAGCCGGGAGCAGC 375

Qy   2283 cgtggccgccagcaaggggctgccgcagcagctgccagcacctccggacgaggatgactc 2342
             |||||||||||||||||||  |||||| | |||||||| |||||||||||||| ||
Db    376 TTCGGCCGCCAGCAAGGGGCTGCAACAGCAGCCGTCAGCACCTTCGGACGAGGATGATTC 435

Qy   2343 agctgccccatccacgttgtccctgctgggccccactttccccggcttaagcagctgctc 2402
           |||||||||||||||| |||||||||||||||||||||||| || ||||||||||||||
Db    436 AGCTGCCCCATCCACGCTGTCCCTGCTGGGCCCCACTTTCCCGGGTTTAAGCAGCTGCTC 495

Qy   2403 cgctgaccttaaagacatcctgagcgaggccagcaccatgcaactccttcagcaacagca 2462
           | | || |||||||||||| |||||||||| |||||||||||| |||||||| |||||
Db    496 CACCGATCTTAAAGACATCTTGAGCGAGGCCGGCACCATGCAACTTCTTCAGCAGCAGCA 555

Qy   2463 gcagga 2468
           ||| |
Db    556 ACAGCA 561
```

```
RESULT   8
AI785019/c
LOCUS       AI785019      637 bp    mRNA            EST       02-JUL-1999
DEFINITION  uj22e09.x1 Sugano mouse kidney mkia Mus musculus cDNA clone
            IMAGE:1920712 3' similar to gb:X53779 Mouse mRNA for androgen
            receptor (MOUSE);, mRNA sequence.
ACCESSION   AI785019
VERSION     AI785019.1  GI:5332735
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 637)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Other_ESTs: uj22e09.y1
            Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:977004
            Seq primer: custom primer used
            High quality sequence stop: 518.
FEATURES             Location/Qualifiers
     source          1. .637
                     /organism="Mus musculus"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1920712"
                     /clone_lib="Sugano mouse kidney mkia"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pME18S-FL3; Site_1: DraIII
                     (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a  DraIII adaptor [TGTTGGCCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert.  Size selection was
                     performed to exclude fragments <1.5kb.  Library
                     constructed by Dr. Sumio Sugano (University of Tokyo
                     Institute of Medical Science).  Custom primers  for
                     sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end
                     primer CGACCTGCAGCTCGAGCACA."
BASE COUNT      211 a    118 c    168 g    139 t      1 others
ORIGIN
```

```
  Query Match             7.0%;  Score 357.8;  DB 10;  Length 637;
  Best Local Similarity   81.4%;  Pred. No. 3.6e-36;
  Matches  529;  Conservative    0;  Mismatches   98;  Indels   23;  Gaps    9;

Qy   4420 cccacatcctgctcaagacgcttctaccagctcaccaagctcctggactccgtgcagcct 4479
          | | |||||||||||  |    |    |||||||||||||   |||   || |||||||||
Db    634 CNCCCATCCTGCTCAGGCGCTTCTACCCAGCTCACCAAG--TCTGATTTCTGTGCAGCCT 577

Qy   4480 attgcgagagagctgcatcagttcacttttgacctgctaatcaagtcacacatggtgagc 4539
          |||||  |||||||||||||||||||||||||| |||||||||||||||| || |||||||||
Db    576 ATTGCAAGAGAGCTGCATCAGTTCACTTTTTACCTGCTAATCAAGTCCCATATGGTGAGC 517

Qy   4540 gtggactttccggaaatgatggcagagatcatctctgtgcaagtgcccaagatcctttct 4599
          |||||||||  | |||||||||||||||||||||||||| |||||||||||||||| ||||
Db    516 GTGGACTTTTCTGAAATGATGGCAGAGATCATCTCTG-GCAAGTGCCCAAGATCCCTTCT 458

Qy   4600 gggaaagtcaagcccatctatttccacacccagtgaagcattggaaa-ccctatttcccc 4658
          ||||||||||||||||||||||||||||||||| ||||||||  |||||||  |||| | |||
Db    457 -GGAAAGTCAAGCCCATCTATTTCCACACACAGTGAAGATTTGGAAACCCCTAATACCCA 399

Qy   4659 accccagctcatgccccctttcagatgtcttctgcctgt--tataactctgcactactcc 4716
          |  ||  |      ||| || ||||||||||||||||||  ||||||||||||||||| |
Db    398 AAACCCACCTTGTTCCCTTTCCAGATGTCTTCTGCCTGTTATATAACTCTGCACTACTTC 339

Qy   4717 tctgcagtgcctt-ggggaatttcctctattgatgtacagtctgtcatgaacatgttcct 4775
          |||||||||||||  ||||||  |||||||| |||||||||||||||||  ||||||  ||||||
Db    338 TCTGCAGTGCCTTGGGGGAAATTCCTCTACTGATGTACAGTCTGTCGTGAACAGGTTCCT 279

Qy   4776 gaattctatttgctgggctttttttttctctttctctccttttcttttttcttcttccctcc 4835
          | |||||||| ||||||||  |  ||            |||| |||||||||||||||||
Db    278 CAGTTCTATTTCCTGGGCTTCTCCTT-----------CTTTTTTTTTCTTCTTCCCTCC 231

Qy   4836 ctatctaaccctcccatggcaccttcagactttgcttcccattgtggctcctatctgtgt 4895
          || | | |||||||||||||||| ||  || | |||| |  ||||||||||||  || |||
Db    230 CTCTTTCACCCTCCCATGGCACATTTTGAATCTGCTGCGTATTGTGGCTCCTGCCTTTGT 171

Qy   4896 tttgaatggtgttgtatgcctttaaatctgtgatgatcctcatatggcccagtgtcaagt 4955
          |||||  |  ||||||||  |||| |||||||||||||||| | |||||||||||||  |
Db    170 TTTGATTTCTGTTGTATTTCTTTGAATCTGTGATGATCCTCTTGTGGCCCAGTGTCAA-T 112

Qy   4956 tgtgcttgtttacagcac--tactctgtgccagccacacaaacgtttacttatcttatgc 5013
          ||||||||||||  |||||  | || ||||||| ||| |||||  |||||||  | |||||||
Db    111 TGTGCTTGTTTATAGCACTGTGCTGTGTGCCAACCAAACAAATGTTTACTCACCTTATGC 52

Qy   5014 cacgggaagtttagagagctaagattatctggggaaatcaaaacaaaaaa 5063
          || || || ||||||||||||  | |||||||  |||   | ||  | | |
Db     51 CATGGCAAATTTAGAGAGCTATAAGTATCTGGAGAAGAAACAAACAGAGA 2
```

```
RESULT   9
AA230070
LOCUS       AA230070      349 bp    mRNA            EST       21-AUG-1997
DEFINITION  nc49b06.s1 NCI_CGAP_Pr3 Homo sapiens cDNA clone IMAGE:1011443
            similar to gb:M23263 ANDROGEN RECEPTOR (HUMAN);, mRNA sequence.
ACCESSION   AA230070
VERSION     AA230070.1  GI:1852363
```

```
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 349)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui, M.D.
            , Michael Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: David B. Krizman, Ph.D.
            cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -41m13 fwd. ET from Amersham
            High quality sequence stop: 230.
FEATURES             Location/Qualifiers
     source          1. .349
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1011443"
                     /clone_lib="NCI_CGAP_Pr3"
                     /sex="Male"
                     /dev_stage="45 years old"
                     /lab_host="DH10B"
                     /note="Vector: pAMP10; Site_1: Not1; Site_2: EcoRI; 1st
                     strand cDNA was primed with oligo(dT)17 on 50 ng of
                     DNAse-treated, total cellular RNA obtained from 5,000-10
                     ,000 microdissected cells histologically-determined to be
                     fully malignant prostate cancer cells.  Double-stranded
                     cDNA was ligated to EcoRI adaptors, 5 cycles of PCR
                     applied to the cDNA with an adaptor-specific primer, and
                     the resulting PCR product subcloned into pAMP10 by the
                     UDG-cloning method (Life Technologies).  Average insert
                     size is 600 bp. NOTE: Not directionally cloned.  This
                     library was constructed by  David Krizman."
BASE COUNT       64 a    105 c     57 g    123 t
ORIGIN
```

```
  Query Match             6.8%;  Score 347.4;  DB 10;  Length 349;
  Best Local Similarity   99.7%;  Pred. No. 8.9e-35;
  Matches  348;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   4540 gtggactttccggaaatgatggcagagatcatctctgtgcaagtgcccaagatcctttct 4599
          ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db      1 GTGGACTTTCCGGAAATGATGGTAGAGATCATCTCTGTGCAAGTGCCCAAGATCCTTTCT 60

Qy   4600 gggaaagtcaagcccatctatttccacacccagtgaagcattggaaaccctatttcccca 4659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGGAAAGTCAAGCCCATCTATTTCCACACCCAGTGAAGCATTGGAAACCCTATTTCCCCA 120

Qy   4660 ccccagctcatgccccctttcagatgtcttctgcctgttataactctgcactactcctct 4719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCCCAGCTCATGCCCCCTTTCAGATGTCTTCTGCCTGTTATAACTCTGCACTACTCCTCT 180

Qy   4720 gcagtgccttggggaatttcctctattgatgtacagtctgtcatgaacatgttcctgaat 4779
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GCAGTGCCTTGGGGAATTTCCTCTATTGATGTACAGTCTGTCATGAACATGTTCCTGAAT 240

Qy   4780 tctatttgctgggctttttttttctctttctctccttttcttttttcttcttcctccctat 4839
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TCTATTTGCTGGGCTTTTTTTTTCTCTTTCTCTCCTTTCTTTTTCTTCTTCCCTCCCTAT 300

Qy   4840 ctaaccctcccatggcaccttcagactttgcttcccattgtggctccta 4888
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CTAACCCTCCCATGGCACCTTCAGACTTTGCTTCCCATTGTGGCTCCTA 349
```

```
RESULT  10
AA659567/c
LOCUS       AA659567      379 bp    mRNA            EST       05-NOV-1997
DEFINITION  nt63a03.s1 NCI_CGAP_Pr3 Homo sapiens cDNA clone IMAGE:1203148
            similar to gb:M23263 ANDROGEN RECEPTOR (HUMAN);, mRNA sequence.
ACCESSION   AA659567
VERSION     AA659567.1  GI:2595721
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 379)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui, M.D.
            , Michael Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: David B. Krizman, Ph.D.
            cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 360.
FEATURES             Location/Qualifiers
     source          1. .379
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1203148"
                     /clone_lib="NCI_CGAP_Pr3"
                     /sex="Male"
                     /dev_stage="45 years old"
                     /lab_host="DH10B"
                     /note="Vector: pAMP10; Site_1: Not1; Site_2: EcoRI; 1st
                     strand cDNA was primed with oligo(dT)17 on 50 ng of
                     DNAse-treated, total cellular RNA obtained from 5,000-10
                     ,000 microdissected cells histologically-determined to be
                     fully malignant prostate cancer cells.  Double-stranded
                     cDNA was ligated to EcoRI adaptors, 5 cycles of PCR
                     applied to the cDNA with an adaptor-specific primer, and
                     the resulting PCR product subcloned into pAMP10 by the
                     UDG-cloning method (Life Technologies).  Average insert
                     size is 600 bp. NOTE: Not directionally cloned.  This
                     library was constructed by  David Krizman."
BASE COUNT      130 a     67 c     95 g     87 t
ORIGIN
```

```
  Query Match             6.8%;  Score 346;  DB 10;  Length 379;
  Best Local Similarity   98.1%;  Pred. No. 1.3e-34;
  Matches  371;  Conservative    0;  Mismatches    5;  Indels    2;  Gaps    2;

Qy   4705 ctgcactactcctctgcagtgccttggggaatttcctctattgatgtacagtctgtcatg 4764
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    376 CTGCACTACTCCTCTGCAGTGCCTTGGGGAATTTCCTCTATTGGTGTACAGTCTGTCATG 317

Qy   4765 aacatgttcctgaattctatttgctgggctttttttttctctttctctcctttctttttc 4824
          |||||||||||||||||||||||||||| |||||| ||||||||||||||||||||||||
Db    316 AACATGTTCCTGAATTCTATTTGCTGGGC-TTTTTTATCTCTTTCTCTCCTTTCTTTTTC 258

Qy   4825 ttcttccctccctatctaaccctcccatggcaccttcagactttgcttcccattgtggct 4884
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    257 CTCTTCCCTCCCTATCTAACCCTCCCATGGCACCTTCAGACTTTGCTTCCCATTGTGGCT 198

Qy   4885 cctatctgtgttttgaatggtgttgtatgcctttaaatctgtgatgatcctcatatggcc 4944
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    197 CCTATCTGTGTTTTGAATGGTGTTGTATGCCTTTAAATCTGTGATGATCCTCATATGGCC 138

Qy   4945 cagtgtcaagttgtgcttgtttacagcactactctgtgccagccacacaaacgtttactt 5004
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    137 CAGTGTCAAGTTGTGCTTGTTTACAGCACTACTCTGTGCCAGCCACACAAACGTTTACTT 78

Qy   5005 atcttatgccacgggaagtttagagagctaagattatctggggaaatcaaaacaaaaaac 5064
          ||||||||||||||||||||||||||||||||||||||||||||||||| | | ||
Db     77 ATCTTATGCCACGGGAAGTTTAGAGAGCTAAGATTATCTGGGGAAATCAAAAC-ACAGAC 19

Qy   5065 aagcaaacaaaaaaaaaa 5082
          ||||||||||||||||||
Db     18 AAGCAAACAAAAAAAAAA 1


RESULT 11
AI315085/c
LOCUS       AI315085      536 bp    mRNA            EST       17-DEC-1998
DEFINITION  uj23g04.x1 Sugano mouse kidney mkia Mus musculus cDNA clone
            IMAGE:1920822 3' similar to gb:X53779 Mouse mRNA for androgen
            receptor (MOUSE);, mRNA sequence.
ACCESSION   AI315085
VERSION     AI315085.1  GI:4030352
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 536)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:977114
            Seq primer: custom primer used
            High quality sequence stop: 487.
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Mus musculus"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1920822"
                     /clone_lib="Sugano mouse kidney mkia"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pME18S-FL3; Site_1: DraIII
                     (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a  DraIII adaptor [TGTTGGCCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert.  Size selection was
                     performed to exclude fragments <1.5kb.  Library
                     constructed by Dr. Sumio Sugano (University of Tokyo
                     Institute of Medical Science).  Custom primers  for
                     sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end
                     primer CGACCTGCAGCTCGAGCACA."
BASE COUNT      181 a    101 c    129 g    124 t      1 others
ORIGIN
```

```
  Query Match             6.3%;  Score 319;  DB 10;  Length 536;
  Best Local Similarity   81.7%;  Pred. No. 2.9e-31;
  Matches  446;  Conservative    0;  Mismatches   81;  Indels   19;  Gaps    6;

Qy   4531 atggtgagcgtggactttccggaaatgatggcagagatcatctctgtgcaagtgcccaag 4590
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    536 ATGGTGAGCGTGGACTTTCCTGAAATGATGGCAGAGATCATCTCTGTGCAAGTGCCCAAG 477

Qy   4591 atcctttctgggaaagtcaagcccatctatttccacacccagtgaagcattggaaaccct 4650
          ||||||||||||||||||||||||||||||||||||||| |||||||| ||||||||
Db    476 ATCCTTTCTGGGAAAGTCAAGCCCATCTATTTCCACACACAGTGAAGATTTGGAAACTAA 417

Qy   4651 atttccccaccccagctcatgccccctttcagatgtcttctgcctgt--tataactctgc 4708
          ||  ||  | |||| ||  |  ||| || ||||||||||||||||  ||||||||||
Db    416 ATACCCANAACCCACCTTGT-TCCCTTTCCAGATGTCTTCTGCCTGTTATATAACTCTGC 358

Qy   4709 actactcctctgcagtgcctt-ggggaatttcctctattgatgtacagtctgtcatgaac 4767
          |||||| |||||||||||||| |||||| |||||||| |||||||||||||||| |||||
Db    357 ACTACTTCTCTGCAGTGCCTTGGGGGAAATTCCTCTACTGATGTACAGTCTGTCGTGAAC 298

Qy   4768 atgttcctgaattctatttgctgggctttttttttctctttctctcctttcttttttcttc 4827
          | |||||| | |||||||| |||||||| |  ||           |||| |||||||||
Db    297 AGGTTCCTCAGTTCTATTTCCTGGGCTTCTCCTT-----------CTTTTTTTTCTTC 250

Qy   4828 ttcctccctatctaaccctcccatggcaccttcagactttgcttcccattgtggctcct 4887
          |||||||||| | | |||||||||||||||| ||  || | |||| |  ||||||||||
Db    249 TTCCCTCCCTCTTTCACCCTCCCATGGCACATTTTGAATCTGCTGCGTATTGTGGCTCCT 190

Qy   4888 atctgtgttttgaatggtgttgtatgcctttaaatctgtgatgatcctcatatggcccag 4947
            || |||||||| |  |||||||  |||| ||||||||||||||| | |||||||
Db    189 GCCTTTGTTTTGATTTCTGTTGTATTTCTTTGAATCTGTGATGATCCTCTTGTGGCCCAG 130

Qy   4948 tgtcaagttgtgcttgtttacagcac--tactctgtgccagccacacaaacgtttactta 5005
          |||||| |||||||||||||| |||||  | || ||||||| |||  |||| |||||||
Db    129 TGTCAA-TTGTGCTTGTTTATAGCACTGTGCTGTGTGCCAACCAAGCAAATGTTTACTCA 71

Qy   5006 tcttatgccacgggaagtttagagagctaagattatctggggaaatcaaaacaaaaaaca 5065
          |||||||||| || || |||| | |||||  | ||||||| |||   | ||  | | | |
Db     70 CCTTATGCCATGGCAAATTTAAAAAGCTATAAGTATCTGGAGAAGAAACAAACAGAGAGA 11

Qy   5066 agcaaa 5071
          |  |||
Db     10 ATAAAA 5


RESULT 12
AZ017963
LOCUS       AZ017963      498 bp    DNA             GSS       25-FEB-2000
DEFINITION  RPCI-23-259K23.TV RPCI-23 Mus musculus genomic clone RPCI-23-259K23
            , DNA sequence.
ACCESSION   AZ017963
VERSION     AZ017963.1  GI:7093347
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 498)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
            ,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-259K23.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
```

```
          Clones are derived from the mouse BAC library RPCI-23. For BAC
          library availability, please contact Pieter de Jong
          (pieter@dejong.med.buffalo.edu). Clones may be purchased from
          BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
          or from Resea ch Genetics (info@resgen.com). BAC end page:
          http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
          Plate: 259  row: K  column: 23
          Seq primer: T7
          Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .498
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-259K23"
                     /clone_lib="RPCI-23"
                     /sex="Female"
                     /lab_host="DH10B"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and  partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI  sites.  The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies).   "
BASE COUNT      102 a    173 c    147 g     75 t      1 others
ORIGIN


  Query Match             5.8%;  Score 293.2;  DB 13;  Length 498;
  Best Local Similarity   73.9%;  Pred. No. 5.3e-28;
  Matches  419;  Conservative    0;  Mismatches   79;  Indels   69;  Gaps    1;

Qy   1865 tggaagtgcagttagggctgggaagggtctaccctcggccgccgtccaagacctaccgag 1924
            |||| |||||||||||||||||||||||||||||||  ||||  | ||||||||||| ||||
Db      1 TGGAGGTGCAGTTAGGGCTGGGAAGGGTCTACCCATGGCCCGCATCCAAGACCTATCGAG 60

Qy   1925 gagctttccagaatctgttccagagcgtgcgcgaagtgatccagaacccgggccccaggc 1984
            |||| |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     61 GAGCGTTCCAGAATCTGTTCCAGAGCGTGCGCGAAGCGATCCAGAACCCGGGCCCCAGGC 120

Qy   1985 acccagaggccgcgagcgcagcacctcccggcgccagtttgctgctgctgcagcagcagc 2044
            |||| |||||||| | |  ||||||||||||||| ||||
Db    121 ACCCTGAGGCCGCTAACATAGCACCTCCCGGCGCCTGTTT---------------------- 160

Qy   2045 agcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2104
                                                        ||||||  ||
Db    161 ----------------------------------------------------ACAGCAGAGGC 171

Qy   2105 agcaagagactagccccaggcagcagcagcagcagggtgaggatggttctccccaag 2164
            || |   |    ||   ||| |  || |||||||||   ||||||||||||||| ||||
Db    172 AGGAGTCNAGCCCCCGGCGGCGGTTGCGGCAGCAGCACACTGAGGATGGTTCTCCTCAAG 231

Qy   2165 cccatcgtagaggccccacaggctacctggtcctggatgaggaacagcaaccttcacagc 2224
            ||||    |||||||||||||||||||||| |||||| |||||||||| ||||||||||
Db    232 CCCACATCAGAGGCCCCACAGGCTACCTGGCCCTGGAGGAGGAACAGCAGCCTTCACAGC 291

Qy   2225 cgcagtcggccctggagtgccaccccgagagaggttgcgtcccagagcctggagccgccg 2284
             |||| | |||   ||| |||||||| |||||  | ||| |||| |||||||| || |||
Db    292 AGCAGGCAGCCTCCGAGGGCCACCCTGAGAGCAGCTGCCTCCCCGAGCCTGGGGCGGCCA 351

Qy   2285 tggccgccagcaaggggctgccgcagcagctgccagcacctccggacgaggatgactcag 2344
            ||  |  |||||||||||||||||||||  ||||| ||||| ||  ||||||||||||
Db    352 CCGCTCCTGGCAAGGGGCTGCCGCAGCAGCCACCAGCTCCTCCAGATCAGGATGACTCAG 411

Qy   2345 ctgccccatccacgttgtccctgctgggccccactttccccggcttaagcagctgctccg 2404
            |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    412 CTGCCCCATCCACGTTGTCCCTGCTGGGCCCCACTTTCCCAGGCTTAAGCAGCTGCTCCG 471

Qy   2405 ctgaccttaaagacatcctgagcgagg 2431
            | ||| ||||||||||  ||| ||||
Db    472 CCGACATTAAAGACATTTTGAACGAGG 498
```

```
RESULT  13
AA229714/c
LOCUS       AA229714      292 bp    mRNA            EST       20-AUG-1997
DEFINITION  nc49b06.r1 NCI_CGAP_Pr3 Homo sapiens cDNA clone IMAGE:1011443
            similar to gb:M23263 ANDROGEN RECEPTOR (HUMAN);, mRNA sequence.
ACCESSION   AA229714
VERSION     AA229714.1  GI:1851877
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 292)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui, M.D.
            , Michael Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: David B. Krizman, Ph.D.
            cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 250.
FEATURES             Location/Qualifiers
     source          1. .292
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1011443"
                     /clone_lib="NCI_CGAP_Pr3"
                     /sex="Male"
                     /dev_stage="45 years old"
                     /lab_host="DH10B"
                     /note="Vector: pAMP10; Site_1: Not1; Site_2: EcoRI; 1st
                     strand cDNA was primed with oligo(dT)17 on 50 ng of
                     DNAse-treated, total cellular RNA obtained from 5,000-10
                     ,000 microdissected cells histologically-determined to be
                     fully malignant prostate cancer cells.  Double-stranded
                     cDNA was ligated to EcoRI adaptors, 5 cycles of PCR
                     applied to the cDNA with an adaptor-specific primer, and
                     the resulting PCR product subcloned into pAMP10 by the
                     UDG-cloning method (Life Technologies).  Average insert
                     size is 600 bp. NOTE: Not directionally cloned.  This
                     library was constructed by  David Krizman."
BASE COUNT      107 a     50 c     73 g     62 t
ORIGIN


  Query Match             5.6%;  Score 285.4;  DB 10;  Length 292;
  Best Local Similarity   99.7%;  Pred. No. 6.1e-27;
  Matches  286;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   4776 gaattctatttgctgggctttttttttctctttctctcctttctttttttcttcttccctcc 4835
            |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    292 GAATTCTATTTGCTGGGCTTTTTCTTTCTCTTTCTCTCCTTTCTTTTTTCTTCTTCCCTCC 233

Qy   4836 ctatctaaccctcccatggcaccttcagactttgcttcccattgtggctcctatctgtgt 4895
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    232 CTATCTAACCCTCCCATGGCACCTTCAGACTTTGCTTCCCATTGTGGCTCCTATCTGTGT 173

Qy   4896 tttgaatggtgttgtatgcctttaaatctgtgatgatcctcatatggcccagtgtcaagt 4955
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    172 TTTGAATGGTGTTGTATGCCTTTAAATCTGTGATGATCCTCATATGGCCCAGTGTCAAGT 113

Qy   4956 tgtgcttgtttacagcactactctgtgccagccacacaaacgtttacttatcttatgcca 5015
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    112 TGTGCTTGTTTACAGCACTACTCTGTGCCAGCCACACAAACGTTTACTTATCTTATGCCA 53
```

```
Qy   5016 cgggaagtttagagagctaagattatctggggaaatcaaaacaaaaa 5062
          ||||||||||||||||||||||||||||||||||||||||||||||
Db     52 CGGGAAGTTTAGAGAGCTAAGATTATCTGGGGAAATCAAAACAAAAA 6
```

RESULT 14
AZ267276

LOCUS	AZ267276 395 bp DNA GSS 26-JUL-2000
DEFINITION	RPCI-23-127M1.TJ RPCI-23 Mus musculus genomic clone RPCI-23-127M1, DNA sequence.
ACCESSION	AZ267276
VERSION	AZ267276.1 GI:9480889
KEYWORDS	GSS.
SOURCE	house mouse.
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 395)
AUTHORS	Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret ,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P. and Fraser,C.M.
TITLE	Mouse BAC End Sequences from Library RPCI-23
JOURNAL	Unpublished (1999)
COMMENT	Other_GSSs: RPCI-23-127M1.TV Contact: Shaying Zhao Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: szhao@tigr.org Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html Plate: 127 row: M column: 1 Seq primer: SP6 Class: BAC ends.

FEATURES	Location/Qualifiers
source	1. .395 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="RPCI-23-127M1" /clone_lib="RPCI-23" /sex="Female" /lab_host="DH10B" /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies). "

BASE COUNT 90 a 103 c 73 g 129 t
ORIGIN

Query Match 4.9%; Score 247.6; DB 13; Length 395;
Best Local Similarity 86.3%; Pred. No. 3.2e-22;
Matches 297; Conservative 0; Mismatches 44; Indels 3; Gaps 2;

```
Qy   4470 cgtgcagcctattgcgagagagctgcatcagttcacttttgacctgctaatcaagtcaca 4529
          | | |  |  ||||| ||||||||||||||||||||||||||||||||||||||||| ||
Db     52 CTTCCTACAGATTGCAAGAGAGCTGCATCAGTTCACTTTTGACCTGCTAATCAAGTCCCA 111

Qy   4530 catggtgagcgtggactttccggaaatgatggcagagatcatctctgtgcaagtgcccaa 4589
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    112 TATGGTGAGCGTGGACTTTCCTGAAATGATGGCAGAGATCATCTCTGTGCAAGTGCCCAA 171

Qy   4590 gatccttttctgggaaagtcaagcccatctatttccacacccagtgaagcattggaaaccc 4649
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||  ||||||||||
Db    172 GATCCTTTCTGGGAAAGTCAAGCCCATCTATTTCCACACACAGTGAAGATTTGGAAACCC 231

Qy   4650 tatttccccaccccagctcatgccccctttcagatgtcttctgcctgt--tataactctg 4707
          || | ||| |  ||  |       ||| || ||||||||||||||||  ||||||||||
Db    232 TAATACCCAAAACCCACCTTGTTCCCTTTCCAGATGTCTTCTGCCTGTTATATAACTCTG 291

Qy   4708 cactactcctctgcagtgcctt-ggggaatttcctctattgatgtacagtctgtcatgaa 4766
          ||||||| |||||||||||||| |||||| |||||||| ||||||||||||||||| ||||
Db    292 CACTACTTCTCTGCAGTGCCTTGGGGGAAATTCCTCTACTGATGTACAGTCTGTCGTGAA 351

Qy   4767 catgttcctgaattctatttgctgggctttttttttctctttct 4810
          || |||||| | |||||||| |||||||  |  ||  | ||| |
Db    352 CAGGTTCCTCAGTTCTATTTCCTGGGCTACTCCTTCTTTTTTTT 395
```

RESULT 15
AI659563/c

LOCUS	AI659563 247 bp mRNA EST 10-MAY-1999
DEFINITION	tu12d12.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:2250839 3' similar to gb:M23263 ANDROGEN RECEPTOR (HUMAN);, mRNA sequence.
ACCESSION	AI659563
VERSION	AI659563.1 GI:4763133
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 247)
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html Seq primer: -40UP from Gibco.

FEATURES	Location/Qualifiers
source	1. .247 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2250839" /clone_lib="NCI_CGAP_Pr28" /sex="male" /dev_stage="adult" /lab_host="DH10B" /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI_CGAP_Pr22 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 985608-986759, 1101192-1101959, and 1217928-1220615). Subtraction by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 78 a 45 c 62 g 62 t
ORIGIN

Query Match 4.7%; Score 238; DB 10; Length 247;
Best Local Similarity 100.0%; Pred. No. 6.1e-21;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy   4825 ttcttccctccctatctaaccctcccatggcaccttcagactttgcttcccattgtggct 4884
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     247 TTCTTCCCTCCCTATCTAACCCTCCCATGGCACCTTCAGACTTTGCTTCCCATTGTGGCT 188

Qy    4885 cctatctgtgttttgaatggtgttgtatgcctttaaatctgtgatgatcctcatatggcc 4944
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     187 CCTATCTGTGTTTTGAATGGTGTTGTATGCCTTTAAATCTGTGATGATCCTCATATGGCC 128

Qy    4945 cagtgtcaagttgtgcttgtttacagcactactctgtgccagccacacaaacgtttactt 5004
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     127 CAGTGTCAAGTTGTGCTTGTTTACAGCACTACTCTGTGCCAGCCACACAAACGTTTACTT 68

Qy    5005 atcttatgccacgggaagtttagagagctaagattatctggggaaatcaaaacaaaaa 5062
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67 ATCTTATGCCACGGGAAGTTTAGAGAGCTAAGATTATCTGGGGAAATCAAAACAAAAA 10
```

Search completed: January 3, 2002, 20:10:32
Job time: 3243 sec